What is personalization worth for Netflix users? Evidence from a text-only and a mock VOD service choice experiment

Bartosz Jusypenko, Michał Paliński

Warsaw University
Faculty of Economic Sciences

Association for Cultural Economics Poland

Personalization vs. privacy: An inevitable trade-off?

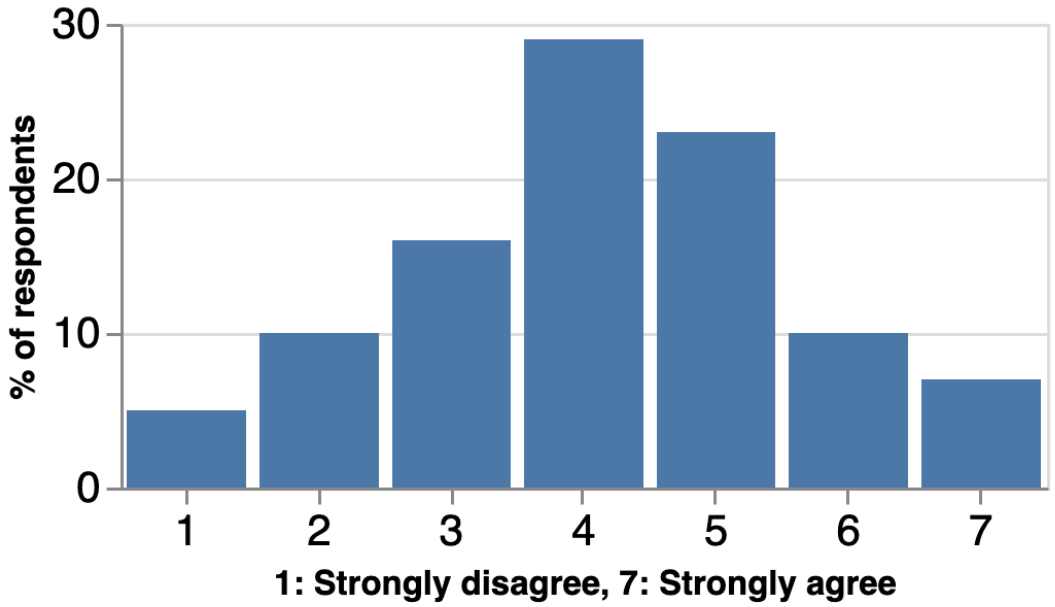
(Awan and Krishnan 2006; Aguirre et al. 2016)

- Status Quo – Netflix has gone from members choosing 2% of the recommended movies to 80% today (Biddle 2021).
- Personalization-privacy paradox:
  - Individuals continue to engage with personalized services despite expressing concerns.
  - Privacy is valued but willingly compromised for benefits such as time and effort savings, improved user experience, and enhanced content discovery.

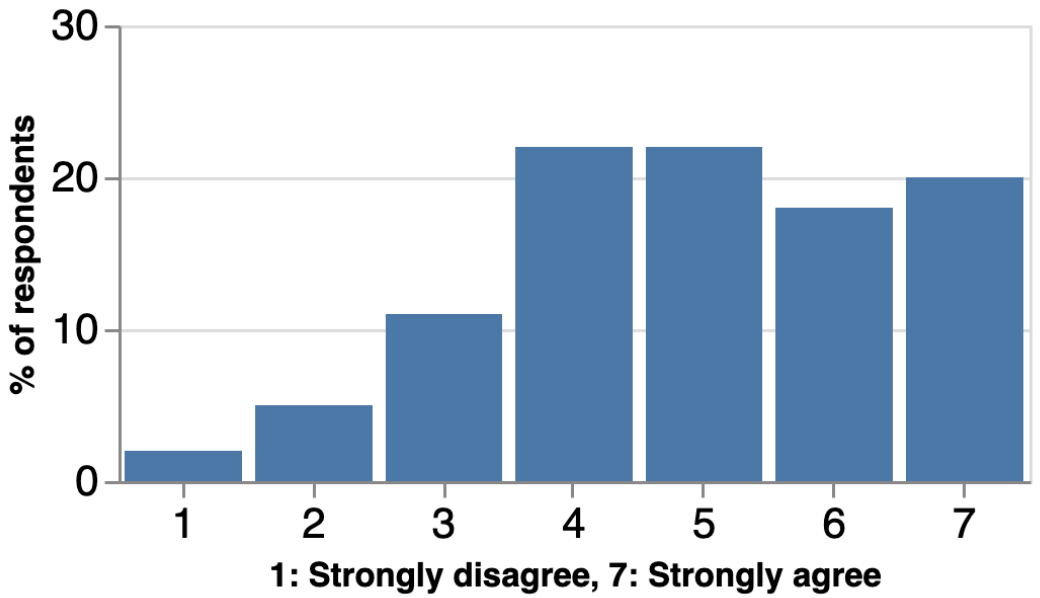
We develop and use our recommender system because we believe that it is core to our business (Netflix 2021).

Stated satisfaction

Netflix recommendations usually align with my taste.



I usually know what I want to watch on Netflix without relying on recommendations.

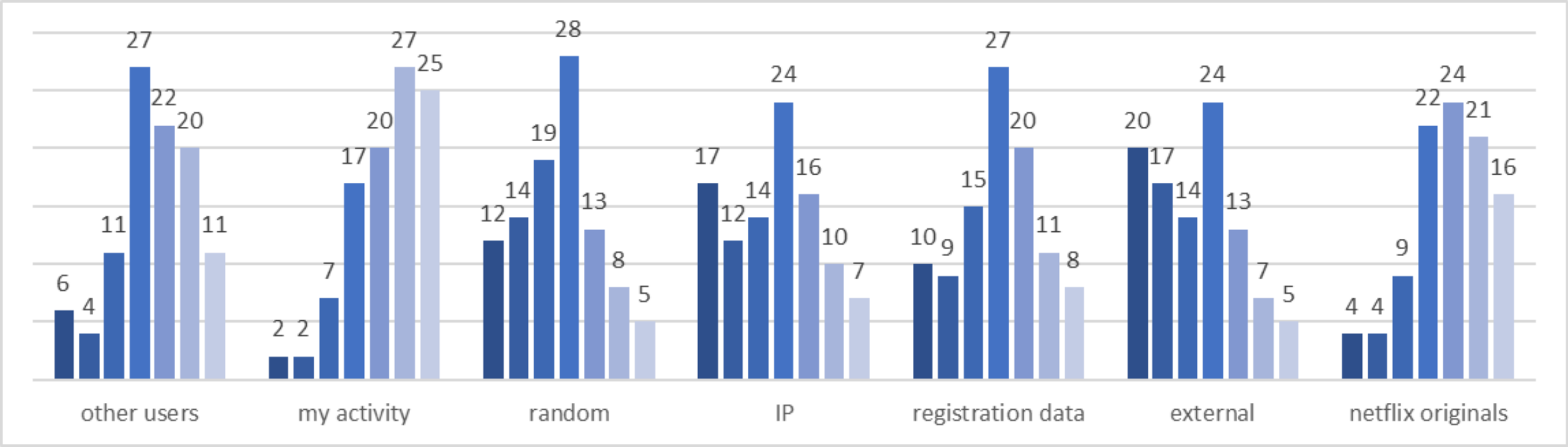


If you could decide, would you share this data with Netflix for personalized recommendations?

- NO:

27% 27% 83% 58% 85%

To what extent do you think the presented content is based on...?



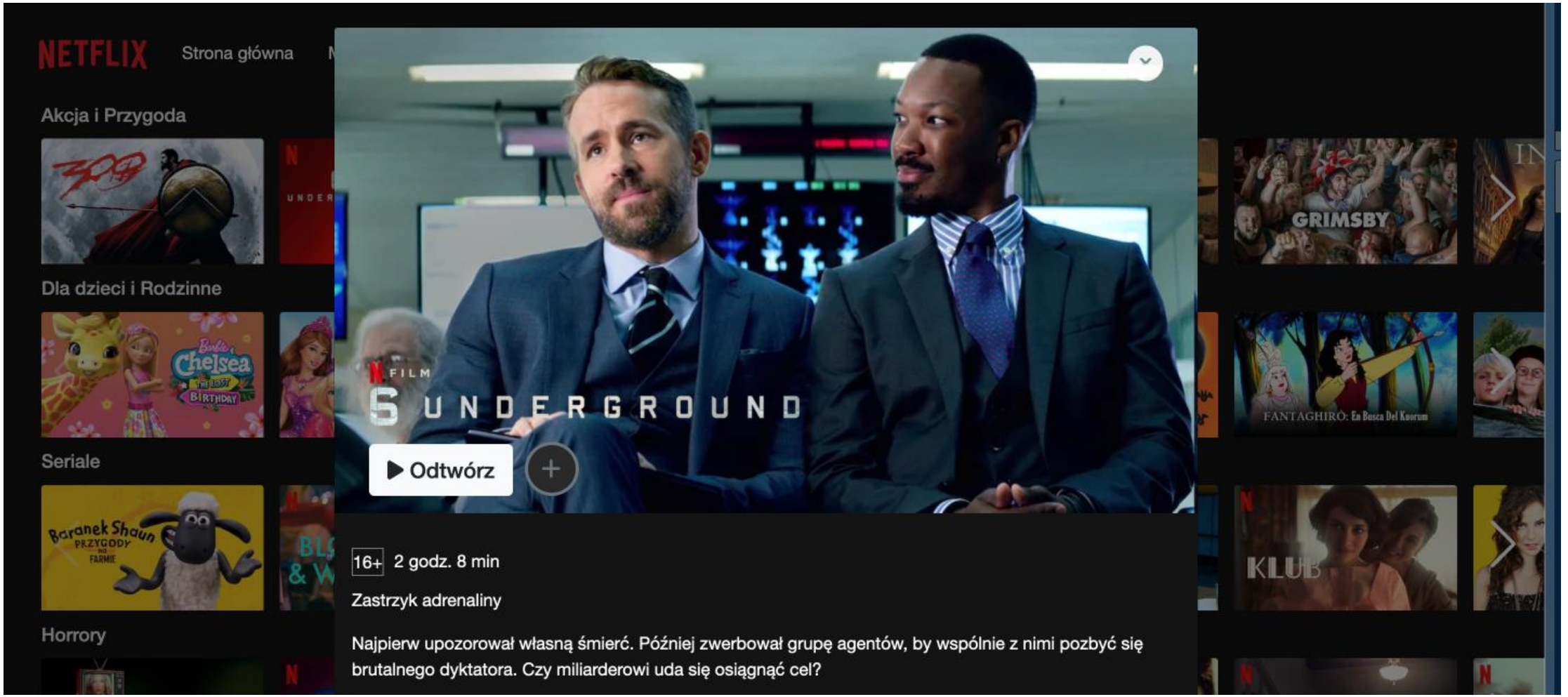
Does Netflix follow you on socials?

- **Information from other sources:** We also obtain information from other sources. We protect this information according to the practices described in this Privacy Statement, plus any additional restrictions imposed by the source of the information. These sources vary over time, but could include:
    - service providers that help us determine a location based on your IP address in order to customize our service and for other uses consistent with this Privacy Statement;
    - security service providers that provide us with information to secure our systems, prevent fraud and help us protect the security of Netflix accounts;
    - payment service providers that provide us with payment or balance information, or updates to that information, based on their relationship with you;
    - publicly-available sources such as publicly available posts on social media platforms and information available through public databases associating IP addresses with internet service providers (ISPs);

Source: https://help.netflix.com/legal/privacy

- Data-driven company? What beyond personalization?

Do you know what would happen if you did not share this data?



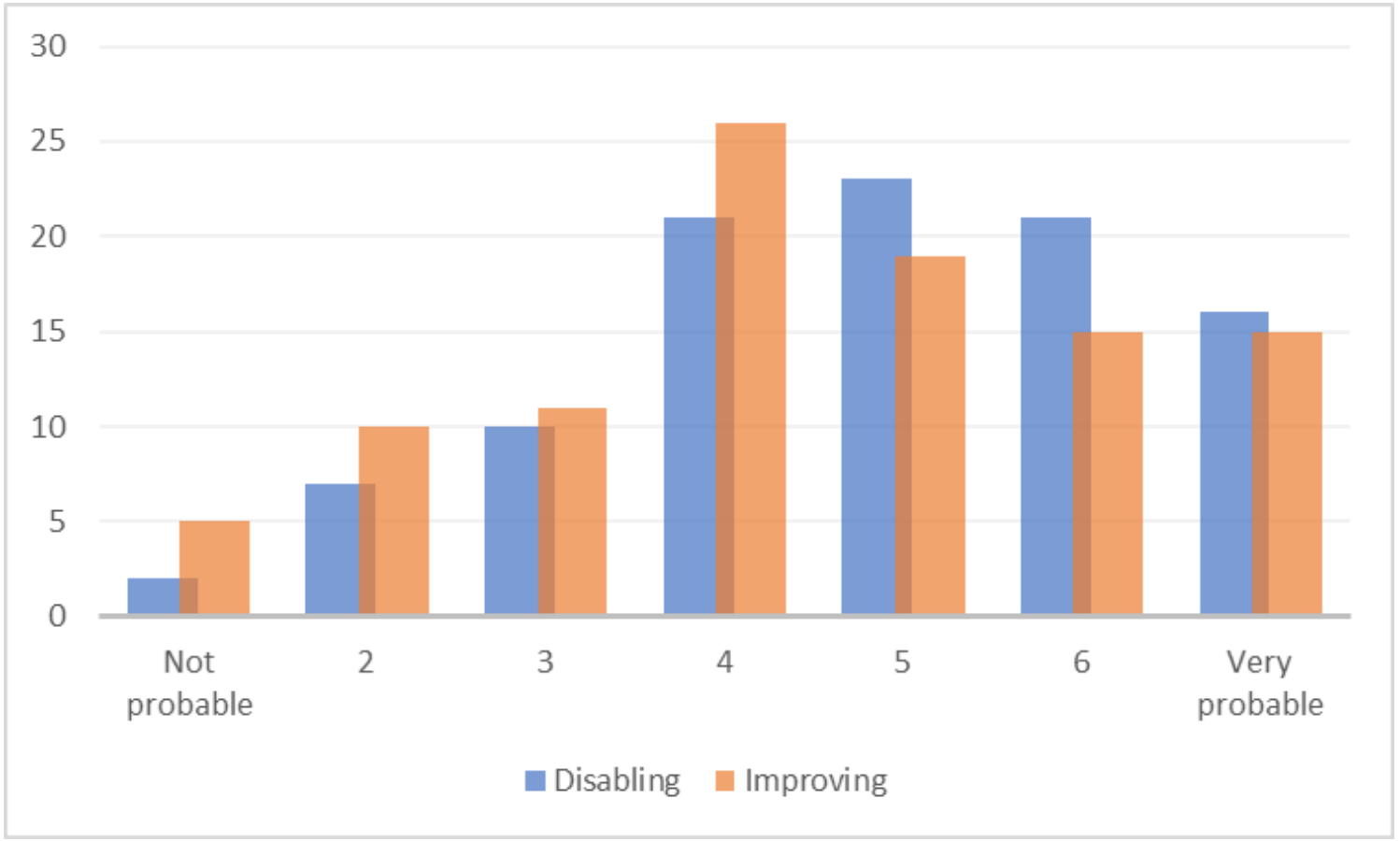
Data collection: choice experiments (2023)

- Choice Experiment 1
  - Data for money?
- Choice Experiment 2
  - Data for convenience?
- Two presentation formats:
  - Text-only experiment;
  - Mock App experiment

- Screening:
  - Netflix user,
  - At least one family member contributes to the subscription.
- Professional polling agency, June 2023, CAWI;
- N=50 (pilot study), N=1003 (text-only DCE), N=216 (mock DCE)
- Representative (gender and age);

Credibility of hypothetical scenarios? New EU regulation to be introduced

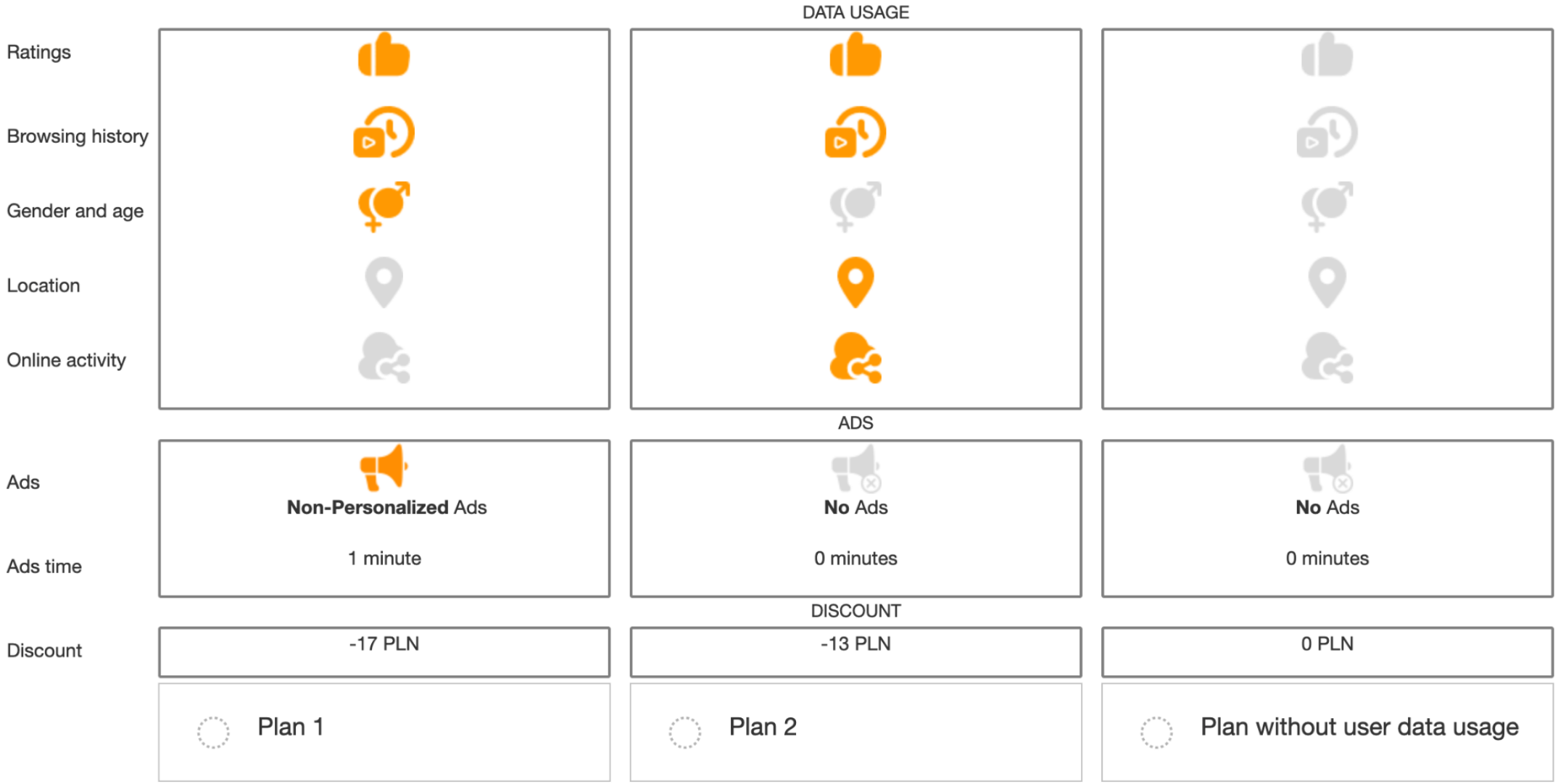
1. Consent needed for the use of data for internal purposes,
2. Consent needed for the use of data for internal purposes, and to provide the service.



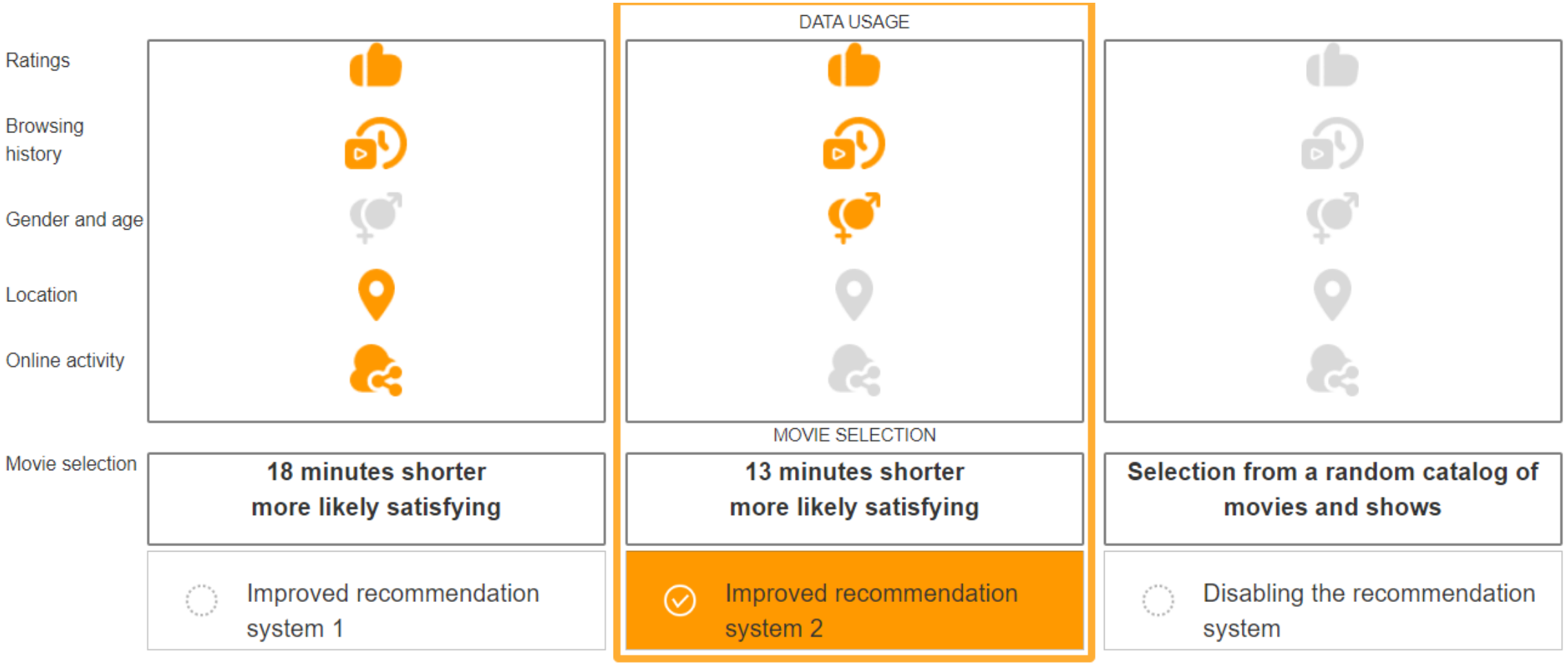
<u>Increased credibility:</u>

- Subscription plans in other countries,
- Ongoing regulatory debate,
- Payment vehicle to reduce protest votes;

Experiment 1: Data for money



Experiment 1: Data for time/convenience



Results: Data for money

- MXL, WTP Space, 1000 Sobol draws;
- 1 PLN = 0.25 USD;

- Substantial preference heterogeneity;
- MU>0 from „internal sharing”
- Very sensitive to the use of data from external sources;
- Against ads, and particularily personalized ads;
- Significant effect of the (experiential) presentation format;

Table 1. Results of the MXL (WTP space, in PLN)

		Mean effect			***St. deviation***		
var.	dist.	coef.		st. err.	coef.		st. err.
Ratings	n	6.448	***	1.423	12.379	***	1.65
Browsing history	n	6.339	***	1.055	-13.371	***	1.129
Gender and age	n	-4.102	***	0.932	-2.271	*	1.697
Location	n	-4.579	***	0.893	-9.749	**	4.131
Socials	n	-10.968	***	1.574	12.542	***	1.952
Pers. Ads	n	-7.117	***	1.429	18.87	***	1.567
Ads Time	n	-1.953	***	0.477	2.651	***	0.5
Discount	n	0.082	***	0.01	-5.146	***	0.01
Mock Netflix effect:							
var.		coef.		st. err.			
Pers. Ads		-5.146	***	3.187			
Ads Time		-0.949	*	0.949			

Note: ***, ** and * indicate 1%, 5% and 10% significance levels, respectively. *n* and *ln* indicate a random parameter distribution (either normal or log-normal).

Results: Data for convenience

- MXL, Preference Space, 500 Sobol draws;
- Time ->
  - *„I am not an average user" effect"*?

- MU>0 from protecting data -> consistent with the previous experiment;
- Substantial preference heterogeneity
  - SD close to mean effects !!!
- V. ambigous effect of the „Mock App" experience;

Classes? ->

Table 1. Results of the MXL (Preference space)

		Mean effect			***St. deviation***		
var.	dist.	coef.		st. err.	coef.		st. err.
ASC (disabling personalization)	n	-2.053	***	0.217	4.566	***	0.303
Gender and age	n	0.563	***	0.06	0.188	*	0.112
Location	n	0.773	***	0.068	-0.41	**	0.176
Socials	n	1.063	***	0.080	-0.924	***	1.952
Time & Convenience	n	-0.053	***	0.014	0.423	***	0.03

Mock Netflix effect:

var.	coef.		st. err.
ASC (disabling personalization)	0.882	*	0.611
Time & Convenience	-0.082	**	0.046

Note: ***, ** and * indicate 1%, 5% and 10% significance levels, respectively. *n* and *ln* indicate a random parameter distribution (either normal or log-normal).

Results: Data for convenience (3-LC)

- LC, Preference Space, 3 classes;

Class 1 – *now is good*
Class 2 – *let's improve*
Class 3 – *you won't get my data!*

Table 1. Results of the 3-class LC model (Preference space)

Mean effects:		***1 CLASS***			***2 CLASS***			***3 CLASS***		
var.	dist.	coef.		st. err.	coef.		st. err.	coef.		st. err.
ASC (disabling personalization)	n	-2.0	***	0.105	-1.434	***	0.426	2.517	***	0.178
Gender and age	n	-0.781	***	0.07	1.01	***	0.214	-2.83	***	0.375
Location	n	-1.053	***	0.072	0.996	***	0.233	-2.299	***	0.343
Socials	n	-1.326	***	0.094	0.945	***	0.233	-3.001	***	0.442
Time & Convenience	n	-0.025	***	0.009	0.175	***	0.03	0.028	***	0.015
Delta					-0.942	***	0.088	-0.295	***	0.067
Mean prob.		0.4685			0.1826			0.3489		

Note: ***, ** and * indicate 1%, 5% and 10% significance levels, respectively.

Conclusions

- Discrepancy between user stated preferences and the perspective presented by Netflix executives.
- Data other than ratings and browsing history remains sensitive.
- Discounts to incentivize users to share more data.
- Users are much more likely to accept random ads than personalized ads.
- Stated preferences may hinge on „experienced disutility”.
- Substantial preference heterogeneity regarding the personalization-privacy trade-off.
- Next steps – one more wave with more intensified „in-app” experience.

Thank you for your attention!

Bartosz Jusypenko
b.jusypenko@uw.edu.pl

Michał Paliński
m.palinski@delab.uw.edu.pl

Warsaw University
Faculty of Economic Sciences

Association for
Cultural
Economics
Poland